DOCUMENT RESUME

ED 109 175 TM 004 622

AUTHOR Scott, Craig S.
TITLE Correlates of Student Ratings of Professorial Performance: Instructor Defined Extenuating Circumstances, Class Size, and Faculty Member's Professional Experience and Willingness to Publish Results.
PUB DATE [Apr 75]
NOTE 18p.; Paper presented at the Annual Meeting of the American Educatial Research Association (Washington, D. C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Rank (Professional); Class Size; *College Students; Correlation; Higher Education; *Predictor Variables; Rating Scales; Statistical Analysis; Student Characteristics; Student Evaluation; *Teacher Characteristics; *Teacher Evaluation; *Teacher Rating

ABSTRACT

This study was undertaken to examine the relationship between the rating professors received from their students and (a) selected professor and student characteristics, and (b) extenuating circumstances frequently expressed by instructors. Data were gathered in the sping of 1974 from an institution-wide administration of a rating scale in 253 classes. Independent variables examined included: class size and composition; instructor's rank, length of service, willingness to publish results; and five extenuating circumstances. Descriptive and correlational procedures were used. The implications of the results for administrative and instructional decision-makers, faculty members, and higher education in general are discussed. (Author/BJG)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED109175

Correlates of Student Ratings of Professorial Performance: Instructor Defined Extenuating Circumstances, Class Size, and Faculty Member's Professional Experience and Willingness to Publish Results

by

Craig S. Scott
Teaching Research Division
Oregon State System of Higher Education

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM004 622

A paper presented at the Annual Meeting of the
American Educational Research Association,
Washington, D. C.: April, 1975

Session 11.06

ABSTRACT

This study examined the relationships between student ratings that college faculty members received from their students and a variety of independent variables. Major independent variables in this study were each faculty member's willingness to publish results of student ratings, experience (as evidenced by rank and highest degree held), size of class (both a 2- and 4-category indicator), and five instructor defined extenuating circumstances. The subjects were 91 faculty members at a small public 4-year college whose academic rank ranged from the level of instructor to full professor. Correlational and descriptive techniques were used in analyzing data from 4,285 separate administrations of a faculty rating scale.

The overall results of the present investigation provide evidence in support of the contention that student ratings of instructor performance are either influenced by or related to a number of variables. Four general conclusions were suggested: (a) there was a slight negative relationship between a faculty member's willingness to publish results of student ratings and the attained student ratings; (b) there was also a slight negative relationship between class size and ratings attained; (c) there was no relationship demonstrated between student ratings and instructor's experience as defined by rank and highest degree held; and (d) there was some indication that there may be a significant relationship between several instructor defined extenuating circumstances and attained student ratings. The first three of these findings have relevance for previously reported research findings. While the findings regarding the extenuating circumstances were not definitive, they do lend support for further examination into the possible influence of such factors on student ratings of instructor performance.

This study is an investigation of the relationships between the ratings given to college faculty members by their students and a variety of independent variables. Some of these variables have been examined elsewhere (i.e., faculty member's rank, highest degree held, and class size) while others have been the subject of little or no research (i.e., faculty member's willingness to publish results of student ratings and five instructor defined extenuating circumstances). In addition to reexamining the previously researched factors, the study represents the first attempt to look at the relationship between student ratings and a new category of extenuating circumstances. This line of inquiry has become more important in light of the movement in higher education toward more extensive use of student ratings in the promotion/tenure process.

In September, 1973, the Teaching Research Division of the Oregon State System of Higher Education (OSSHE) began a 3-year study of the factors which influence promotion and tenure in higher education. The study is supported by a grant from the Fund for the Improvement of Postsecondary Education (HEW) and is conducted as a part of the Division's Higher Education Research Program. The basic intent of this larger study is to develop a higher degree of understanding of the consequences resulting from present policies and procedures surrounding the assessment of faculty performance, while making attempts to assist OSSHE institutions in making any necessary refinements in existing practices. To date, a number of data sources have been used. A primary source has been administrators at the institution and department levels. These administrators supplied policy statements, guidelines, forms, and indicators used when they assess faculty performance. The second major source of data was a system-wide survey which evaluated the

promotion/tenure process from faculty members' perspectives. Finally, many of the campuses used faculty review committees at the department level and then at intermediate levels on up to the president of the institution. Where such campus-wide personnel committees or advisory bodies existed, project staff have interviewed these individuals, raising a number of questions about those features of the promotion/tenure process which they find the most useful and about those features which they deem to be in considerable need of change.

During the course of these system-wide data collection efforts, both teaching and administrative faculty voiced a number of concerns about the ramifications of the use of student rating data in the promotion/tenure process. It soon became evident that many of the expressed concerns were a direct result of a general lack of familiarity with available research which examines the relationships between various factors and student ratings. However, a number of faculty members were concerned about the possible influence of a number of extenuating circumstances on student ratings which they felt could be of legitimate concern to a faculty member whose ratings may be used in the promotion/tenure process. The present study constitutes an attempt to begin to clarify the degree of legitimacy surrounding these extenuating circumstances. The results of the investigation should be of interest to teaching and administrative faculty, regardless of whether or not the extenuating circumstances tend to result in significantly different student ratings. The important point here is that those who are charged with the responsibility of promotion and/or tenure decisions should know which, if any, of the extenuating circumstances appear to have impact upon student ratings and are therefore legitimate concerns, and which are not.

Previous research (Heilman & Armentrout, 1936; Lovell & Hanner, 1955; McDaniel & Feldhusen, 1971) has yielded results which suggest that there may be a slight negative correlation between class size and student ratings. Guthrie (1954) concluded that there was no correlation between size of class and ratings, while other studies have indicated that there may be a variable relationship (Gage, 1961; Wood, Linsky, & Straus, 1974).

With regard to the association between a faculty member's experience (usually represented by rank) and student ratings, the relationship is unclear. Langen (1966) reported a positive correlation between rank and ratings. Conversely, a study by Riley, Ryan, & Lifschitz (1950) suggested a negative correlation between rank and ratings. The present study examines the relationshipship between faculty members' experience and student ratings using two indicators (viz., faculty member's rank and faculty member's highest degree held).

Method

Subjects

The subjects were 91 faculty members at a small public 4-year college whose academic rank ranged from the level of instructor to full professor (4 instructors, 46 assistant professors, 21 associate professors, and 20 full professors). All subjects were voluntary participants in a campus-wide administration of a student rating scale. Approximately 60% of the faculty members at the institution participated in the evaluation. Data were collected in May, 1974, in 253 classes during the seventh week of instruction.

Instruments

During the winter quarter of the 1973-74 academic year, the institution's Faculty Senate established a Joint Faculty Evaluation Committee. This

committee was composed of six faculty members and six students. The author, who was not a member of the faculty, was asked to serve as a consultant to the committee. Consultation was provided in the areas of faculty rating scale selection, administration, and analysis. The primary objective of this committee was to devise a complete student rating system for annual, campus-wide use. Subsequent to examination of a variety of existing student rating scales, the committee decided that none fully met the needs of faculty evaluation at their specific institution. The committee decided to develop its own rating scale by combining items from three rating scales.

The resultant rating scale consisted of four sections. The first section contained 13 items intended for campus-wide use (12 specific items and 1 overall instructor performance item). Prior to selection, each potential item was examined in an attempt to eliminate those items which might result in an unfair advantage or disadvantage for a faculty member because of his major discipline and/or course assignment. In addition, each item was designed to avoid requesting any information from the student that would require judgments about the feelings or attitudes of classmates. Each of the 12 specific performance items was tied to a Likert-type 5-point scale, which ranged from a low of "1" (strongly disagree) to a high of "5" (strongly agree). The second and third section of the total rating scale were optional department and instructor segments, respectively. A final section provided an avenue students could use to supply written comments directly to their instructors. Only the institution-wide items, constituting section 1, were analyzed in this study.

The content of the 13 items was as follows: "the instructor seems to be knowledgeable in the subject area with which the course deals"; "the

instructor has effectively presented the subject matter to be learned (in this evaluation, 'subject matter' may mean facts, skills, insights, and/or attitudes)"; "the instructor seems to be genuinely interested in the subject matter"; "the instructor seems to be genuinely interested in the students' learning the material"; "the instructor is someone I would feel comfortable approaching to ask questions or to discuss ideas relating to the subject matter"; "the instructor has been reasonably accessible to me outside of class (if you did not try to contact the instructor, leave this item blank)"; "the instructor seems to be sensitive to the responses of the class"; "the course seems to be sufficiently rigorous (challenging)"; "the instructor has presented material which I think is relevant (within the limitations imposed by the subject)"; "I have learned a great deal in this course"; "my interest in the subject matter has increased as a result of taking this course"; and "the instructor has been fair in the selection of examination questions and/or evaluation criteria used in this course." The overall (general) rating item was, "Rank the instructor on the following scale: 1 = unacceptable, 2 = below standard, 3 = standard, 4 = above standard, and 5 = outstanding."

An Instructor Data Sheet was used to collect information from the faculty members about both class-related characteristics and instructor defined extenuating circumstances.

Procedure

Approximately ten days before the week of the evaluation, all faculty members at the institution received a copy of a memorandum from the committee which asked if they were willing to participate in the evaluation, and if they were willing to release the resulting data for publication by the student

body government. Each participating faculty member was requested to identify a student administrator for each of his classes. These students were responsible for the actual administration of the rating scale. Several student administrator orientation sessions were held the week before the evaluation for the purpose of familiarizing them with the process they were to follow when administering the rating scale in their classes. Detailed administration instructions were provided orally at these sessions. The student administrators also received written administration instructions, at the time they picked up the rating scale packets, immediately prior to the time of the evaluation in their respective classes.

Each rating scale packet contained sufficient copies of the rating scale, machine processable answer sheets, blank sheets of paper for written comments, and pencils. In addition, each packet contained written administration directions for the student administrators and an Instructor Data Sheet. Each returned packet was hand checked by the author to verify that the students had correctly coded their course identification information and evaluative responses (students responded anonymously--however, 95% also answered five descriptive items which provided information about their personal characteristics). The percentage distribution of these key background variables (i.e., class standing, overall GPA, required or elective course, age, and sex) were compared with information published by the Registrar's Office on the same variables. The distribution corresponded quite closely between the two data sources for each variable. All of these data were machine scored and merged with information provided by the appropriate faculty member about each class. These combined data provided input for both the individual instructor reports, which were provided each participant, and for the analyses of this study.

Variables

The primary dependent variable used in this study was the mean student rating on the total scale (items 1 through 12). Also examined as a dependent variable was the mean rating received by the faculty member on the overall (general) item.

Independent variables came from a variety of sources. Of the 10 independent variables included in this study, 4 came from institutional records and 6 were supplied by participating faculty members. Institutional records provided data on class size (a 4-category indicator and a 2-category indicator were constructed) and on faculty members' highest degree and rank. Willingness to publish results of the ratings was ascertained from the original participation inquiry. Data on instructor defined extenuating circumstances was gathered by means of an item on the Instructor Data Sheet (a separate sheet was completed for each evaluated class). The item read as follows: "An instructor's performance can sometimes be hindered by circumstance. If you feel there are current circumstances which may unfairly lower the ratings given you by your students, please indicate them." The faculty member was instructed to check "Yes" or "No" with regard to five possible extenuating circumstances. These circumstances included: "Work load too heavy to devote adequate attention to the course," "class is too large to adequately present the material," "this particular class is outside of my greatest competence area," "this is a new/untried course," and "innovations in the instruction are being introduced." Space was provided to add "other" extenuating circumstances, but none were suggested.

Analysis

Correlational and descriptive techniques were used in analyzing three different data sets. The first data set contained responses from 4,285

separate administrations of the faculty rating scale. This data set served two purposes. First, it was used to calculate item analysis information related specifically to the rating scale. Second, it was used to create two other summary data sets--one for courses and one for instructors. The course data set contained data from 253 courses (prior to analysis 16 courses were excluded due to special circumstances). The instructor data set contained one randomly selected course for each of the 91 participating faculty members. These summary data sets consisted of: (a) information from course and/or instructor summaries of the 4,285 student responses, and (b) course-specific information supplied by faculty members as previously described. The course data set (237 courses) was used to examine, in a correlational sense, the relationship between class size and student ratings. The instructor data set was used to examine relationships between independent variables related to instructors, such as, rank, highest degree held, willingness to publish results of student ratings, and the five previously described extenuating circumstance items.

Results

Major independent variables in this study were each faculty member's willingness to publish results of student ratings, experience (as evidenced by rank and highest degree held), size of class (both a 2- and 4-category indicator), and five instructor defined extenuating circumstances. Willingness to publish, instructor experience, and size of class were analyzed separately from the extenuating circumstances.

Willingness to Publish, Instructor Experience, and Size of Class

Table 1 presents correlations between the dependent and independent variables using data from the Spring, 1974, student evaluations of

Table 1

Correlations between Selected Variables Using Data from the Spring, 1974, Student Evaluations of Instructor Performance[1]

Variable Description	Row Variable Number						
	1	2	3	4	5	6	7
1. Mean rating on total scale (12 items)	——	.92*	.24*	-.15	-.08	-.37*	-.32*
2. Mean rating on an overall (general) item	.89*	——	.25*	-.13	-.04	-.34*	-.27*
3. Faculty member's willingness to publish results of ratings (1=no; 2=maybe; 3=yes)	.13*	.13*	——	-.27*	.00	.01	.01
4. Faculty member's rank (1=instructor; 2=asst. prof.; 3=assoc. prof.; 4=prof.)	-.17	-.12	-.14*	——	.69*	-.07	-.08
5. Faculty member's highest degree (1=bachelor's; 2=master's; 3=Ph.D.)	-.18	-.11	.04	.70*	——	.08	.11
6. Class size A (1=1 to 10; 2=11 to 30; 3=31 to 50; 4=over 60)	-.23*	-.23*	.00	.07	.11	——	.87*
7. Class size B (1=30 or less; 2=over 30)	-.24*	-.21*	-.03	.08	.11	.84*	——

[1] Correlations above the diagonal are based upon 91 classes (one class for each participating faculty member with 7 classes excluded due to special circumstances). Correlations below the diagonal are based upon 237 classes (one to four classes for each participating faculty member).

* Significant beyond the .05 level.

instructor performance. Correlations above the diagonal were based upon 91 classes (one class for each participating faculty member with 7 classes excluded due to special circumstances). Correlations below the diagonal were based upon 237 classes (one to four classes for each participating faculty member). Resultant correlations were based first on one class for each participating faculty member and second on multiple classes for each participating faculty member, because three of the variables (i.e., willingness to publish results of ratings, faculty member's rank, and faculty member's highest degree held) would have been confounded if analyzed in a situation in which faculty members were represented by varying numbers of

classes. Likewise, the two class size indicators were more appropriately examined using data which were based on more than one class per faculty member.

Examination of Table 1 indicates a strong relationship between the mean ratings on the total scale (12 items) and the mean rating on the overall instructor rating item (.89 for the multiple class data set vs. .92 for the single class data set). This implies that either score could have been validly used as a criterion measure against which to evaluate the relationship between the independent variables of interest and the dependent variable (student ratings). Note that in both data sets each of the independent variable's two correlations with the two dependent variables exhibited extremely slight variations. Therefore, for the sake of simplicity the following discussions will focus primarily on the relationships between the independent variables and the mean rating on the total scale.

Willingness to publish results. Examination of Table 1 shows that there was a correlation of .24 between faculty member's willingness to publish results of student ratings and the mean rating on the total rating scale. This suggests that faculty members who, for one reason or another, did not release their data for publication tended to get slightly lower ratings than did those instructors who were willing to have the student government publish the results of their ratings. (It is important to note that permission to publish was obtained prior to the week of the evaluations.)

Instructor experience. The professorial rank which the faculty members held and their highest academic degree were used in this study as indicators of professional experience. Examination of Table 1 indicates that neither of these indicators were significantly related to attained student ratings.

Correlations for faculty member's rank and highest degree held were -.15 and -.08, respectively.

Size of class. Table 1 shows significant negative correlations of -.24 and -.23 for the 2- and 4-category class size indicators. This finding indicates that for this sample of faculty members, there was a slight tendency for student ratings to decrease as class size increased.

Instructor Defined Extenuating Circumstances

Table 2 presents mean ratings and t-scores for five instructor defined extenuating circumstances. Since most faculty members reported that the extenuating circumstances did not apply, there was a substantial difference in the sizes of the group responding "yes" and the group responding "no." For this reason it was decided that a correlational analysis would be inappropriate because it would tend to obscure any existing relationship due primarily to this discrepancy in group size. Therefore, t was used as a test statistic to test the hypothesis of no difference between the means (Blommers & Lindquist, 1960).

Table 2

Mean Ratings and t-scores for Instructor Defined Extenuating Circumstances

Instructor defined extenuating circumstance	Mean rating of group responding "No"	Mean rating of group responding "Yes"	t of difference between means
1. "Work load too heavy to devote adequate attention to the course"	4.2	4.0	-1.52 ($p < .2$)*
2. "Class too large to adequately present the material"	4.1	3.7	-1.74 ($p < .1$)
3. "This particular class is outside of my greatest competence area"	4.1	4.3	.68 ($p < .5$)
4. "This is a new/untried course"	4.1	4.0	-1.07 ($p < .3$)
5. "Innovations in the instruction are being introduced"	4.2	4.1	- .61 ($p < .6$)

* p is the probability of having t this large or larger in size by chance.

Examination of the group means in Table 2 indicates that there was a tendency for student ratings to decrease for those instructors who said that the extenuating circumstances applied to their particular situation. Examination of the associated t-scores suggests that two of the extenuating circumstances produced group mean differences significant beyond the .2 level. The remaining three extenuating circumstances provide less evidence of a significant relationship between student ratings and these extenuating circumstances. While these results are not definitive, they do lend support to further examination into the possible influence of such factors on student ratings of instructor performance.

Discussion

The overall results of the present investigation provide evidence in support of the contention that student ratings of instructor performance are either influenced by or related to a number of variables. Four general conclusions were suggested: (a) there was a slight negative relationship between a faculty member's willingness to publish results of student ratings and the attained student ratings; (b) there was also a slight negative relationship between class size and ratings attained; (c) there was no relationship demonstrated between student ratings and instructor's experience as defined by rank and highest degree held; and (d) there was some indication that there may be a significant relationship between several instructor defined extenuating circumstances and attained student ratings. The first three of these findings have relevance for previously reported research findings. With regard to the fourth conclusion, the present study was not expected to find conclusive evidence that all, or even any, of the

instructor defined extenuating circumstances were significantly related to student ratings of instructor performance. Whether or not the relationship between the extenuating circumstances and student ratings is large enough to be of real practical significance to faculty members, is a question about which more data are needed before any reliable answers can be given.

The proposition that students have the most advantageous perspective, by virtue of their daily classroom observations, to assess the in-class effectiveness of their professors is the topic of a diminishing amount of debate. Comprehensive evaluation of teaching performance for either salary and promotion considerations or for instructional improvement has become more and more difficult to achieve and justify unless some type of data from students is available. However, as ratings become more influential, it becomes even more important to understand what it is that enables an instructor or a group of instructors, to attain good ratings.

Unfortunately, there has been little research completed which deals directly with administrative uses of student rating data. While the present report also does not directly involve data emanating from or leading to administrative use of ratings, the results of the extenuating circumstance analyses are unique and should be of direct interest to both individual faculty members who feel that their performance in the classroom can sometimes be hindered by circumstance, and to administrators who must interpret the results of student ratings.

The fact that the focus of this study is student ratings should not imply that such ratings are viewed by the author as being the sole, or even the most justifiable, indicator of teaching performance in an overall decision-making scheme. It is recognized that the total performance of an instructor is a function of a variety of very different types of behaviors.

Performance in the classroom per se is merely part of a professor's role that can be reflected in student ratings. Data which relate directly to faculty members' classroom performance (i.e., student ratings) have a necessary and justifiable role to fulfill in the overall assessment of professorial performance. However, student ratings should only be used in conjunction with other sources of information about instructional effectiveness.

An argument can be made that this study, because it is the first large-scale look at these instructor defined extenuating circumstances, is not truly representative of the actual relationship between the extenuating circumstances and student ratings. This seems to be a reasonable argument, particularly in light of some of the cited limitations inherent within the data which, without question, limited the scope of the analyses. Therefore, it would be reasonable to continue this type of investigation using more precise definitions of extenuating circumstances in conjunction with a rating scale about which faculty are more familiar than they were in the present case. In addition, utilization of other techniques such as analysis of variance would add a needed dimension to the investigation that would allow much more to be said about the legitimacy of the extenuating circumstances.

In a time of growing use and possible misuses of student ratings for promotion and tenure granting purposes, the identification of relevant extenuating circumstances which may unfairly influence student ratings, constitutes a step toward a much needed refinement in the process of evaluating this aspect of faculty members' teaching performance. However, in light of the present findings, a large part of the task remains if we are to determine which, if any, of the often-voiced instructor defined

extenuating circumstances are legitimate concerns, and which are not. The major point here is that it is imperative that further investigation into the possible influence of extenuating circumstances, such as those examined in this study, be continued. Until the influence of such factors is either confirmed or discounted, it will remain difficult to equitably apply the results of student ratings in decision-making processes.

References

Blommers, P., & Lindquist, E. F. Elementary statistical methods in psychology and education. Boston: Houghton Mifflin, 1960.

Gage, N. L. The appraisal of college teaching. Journal of Higher Education, 1961, 32, 17-22.

Guthrie, E. R. The evaluation of teaching: A progress report. Seattle: University of Washington, 1954.

Heilman, J. D., & Armentrout, W. D. The rating of college teachers on ten traits by their students. Journal of Educational Psychology, 1936, 27, 197-216.

Langen, T. D. F. Student assessment of teaching effectiveness. Improving College and University Teaching, 1966, 14, 22-25.

Lovell, G. D., & Hanner, C. F. Forced-choice applied to college faculty rating. Educational and Psychological Measurement, 1955, 15, 291-304.

McDaniel, E., & Feldhusen, J. F. College teaching effectiveness. Today's Education, 1971, 60, 27.

Riley, J. W., Ryan, B. F., & Lifschitz, M. The student looks at his teacher. New Brunswick, N. J.: Rutgers University Press, 1950.

Wood, K., Linsky, A. S., & Straus, M. A. Class size and student evaluations of faculty. Journal of Higher Education, 1974, 45, 524-534.